Title: Customer Sales Analysis

By

Suhas S

(2348563)

Name of the Programme: Master of Science (Artificial Intelligence & Machine Learning)

Tableau Dashboard Analysis

Date of Submission: 19-01-2024

Contents

Visualization Platform:

Tableau:

Tableau is a powerful and intuitive data visualization platform that empowers users to convert raw data into meaningful insights. Designed for both beginners and seasoned analysts, Tableau facilitates the exploration and analysis of complex datasets through interactive and dynamic visualizations. Its user-friendly interface allows for seamless drag-and-drop functionality, enabling users to create compelling charts, graphs, and dashboards without the need for extensive coding or technical expertise.

1. Dataset Overview:

Introduction:
This study delves into a comprehensive review of a sales dataset, encompassing critical attributes such as order details, item specifics, and customer information. The primary objective is to analyze the dataset's structure, identify key trends, and propose potential visualizations for an effective Tableau report.

Dataset Overview:
The dataset comprises a multitude of fields, including order_id, order_date, item_id, price, category, and various customer-related attributes. A thorough understanding of each attribute is imperative for the creation of meaningful visualizations in Tableau.

Data Quality and Completeness:
Prior to commencing the analysis, a rigorous assessment of data quality is essential. This involves identifying and addressing missing values, outliers, and inconsistencies, ensuring the dataset is conducive to accurate and reliable visualizations within Tableau.

Temporal Analysis:
Leveraging Tableau's temporal analysis capabilities, a scrutiny of trends over time will be conducted. Visualizations such as line charts or heatmaps will be employed to decipher patterns, seasonality, and potential areas for improvement in metrics such as sales value, quantity ordered, and discount amounts.

Sales Performance by Category:
Categorization of sales data will be pivotal in understanding the performance of different product categories. Visual representations, such as bar charts or tree maps, will be explored to effectively communicate which categories contribute most significantly to overall sales.

Customer Segmentation:
Utilizing demographic information, visualizations will be crafted to segment customers based on attributes such as age, gender, and location. This segmentation will facilitate targeted marketing strategies and personalized customer engagement.

Payment Method Analysis:
A detailed analysis of preferred payment methods will be undertaken. Visualizations such as pie charts or bar graphs will be employed to illustrate the distribution of payment methods, providing insights for strategic decisions related to payment processing.

Discount Impact on Sales:
The relationship between discounts and sales will be investigated. Visual aids such as scatter plots or regression analyses will be utilized to discern whether discounts significantly impact the quantity ordered or total sales value, informing optimal pricing strategies.

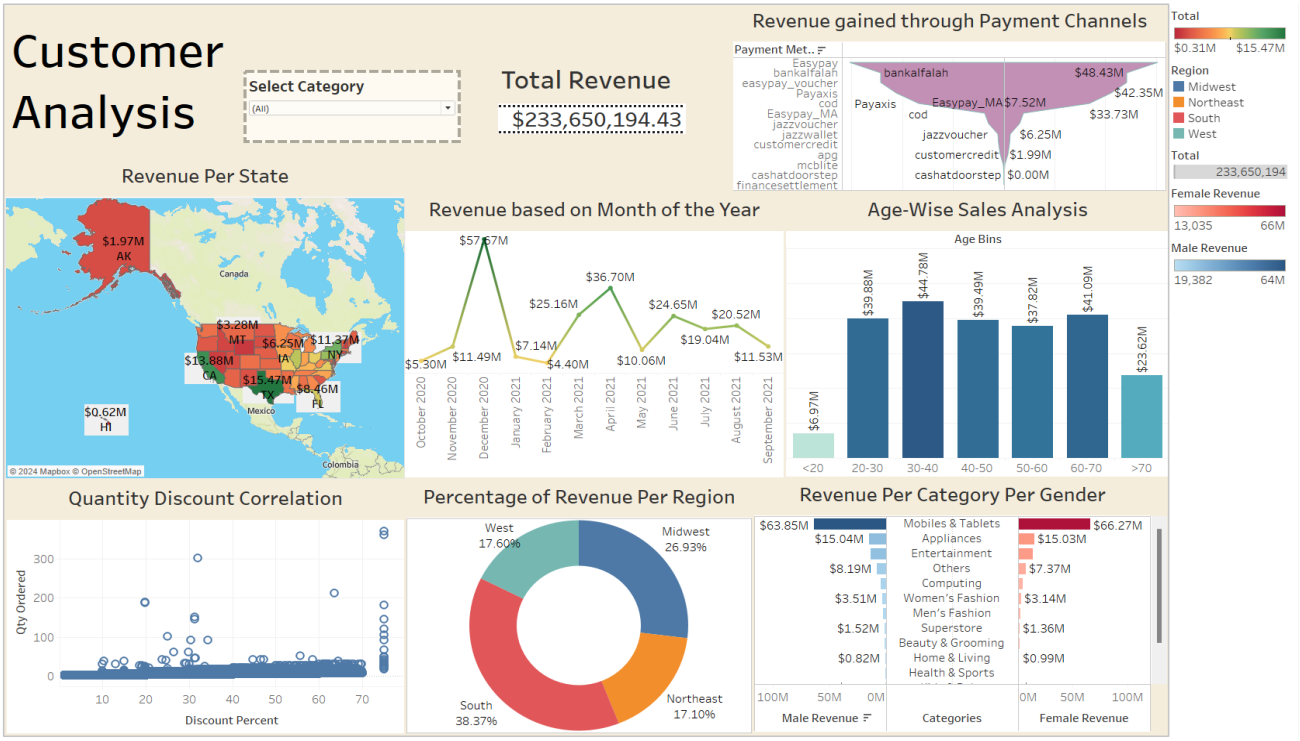
Geospatial Analysis:
Geographic information, including city, state, and zip code, will be utilized for geospatial analysis. Map visualizations in Tableau will be employed to illustrate regional sales patterns, aiding in the identification of growth areas and regions requiring additional attention.

Customer Retention and Acquisition:
A focused analysis of customer retention rates over time will be conducted. Visualizations such as cohort analysis or line charts will provide insights into customer loyalty, facilitating the development of tailored retention strategies.

Conclusion:
In conclusion, this comprehensive review outlines key areas for exploration within the sales dataset, employing Tableau's robust visualization capabilities. The objective is to transform raw data into actionable insights, fostering informed decision-making and strategic planning within the context of the envisioned Tableau report. The narrative flow and coherence of each

visualization should be carefully considered to present a compelling and informative story.

2. Dashboard Description:



The dashboard provides a comprehensive overview of revenue gained through payment channels for a company. It includes various charts and graphs to visualize the data from multiple perspectives.

Key metrics:

Total revenue: This is the overall revenue generated by the company, displayed at the top of the dashboard.

Revenue by payment channel: This is shown in the funnel chart at the top right. The chart shows that "Easypay" is the most popular payment channel, followed by "bankalfalah" and "Payaxis".

Revenue by region: This is shown in the pie chart at the top right. The chart shows that the "south" region generates the most revenue, followed by the "MidWest" and "West" regions.

Revenue by month: This is shown in the line graph at the middle. The graph shows that revenue generated for period of 1 Year.

Age-wise sales analysis: This is shown in the bar chart at middle right. The chart shows the revenue generation distribution across various age groups.

Revenue per category per gender: This is shown in the bar chart at the bottom right. The chart shows that the "Mobiles & Tablets" category generates the most revenue for female customers.

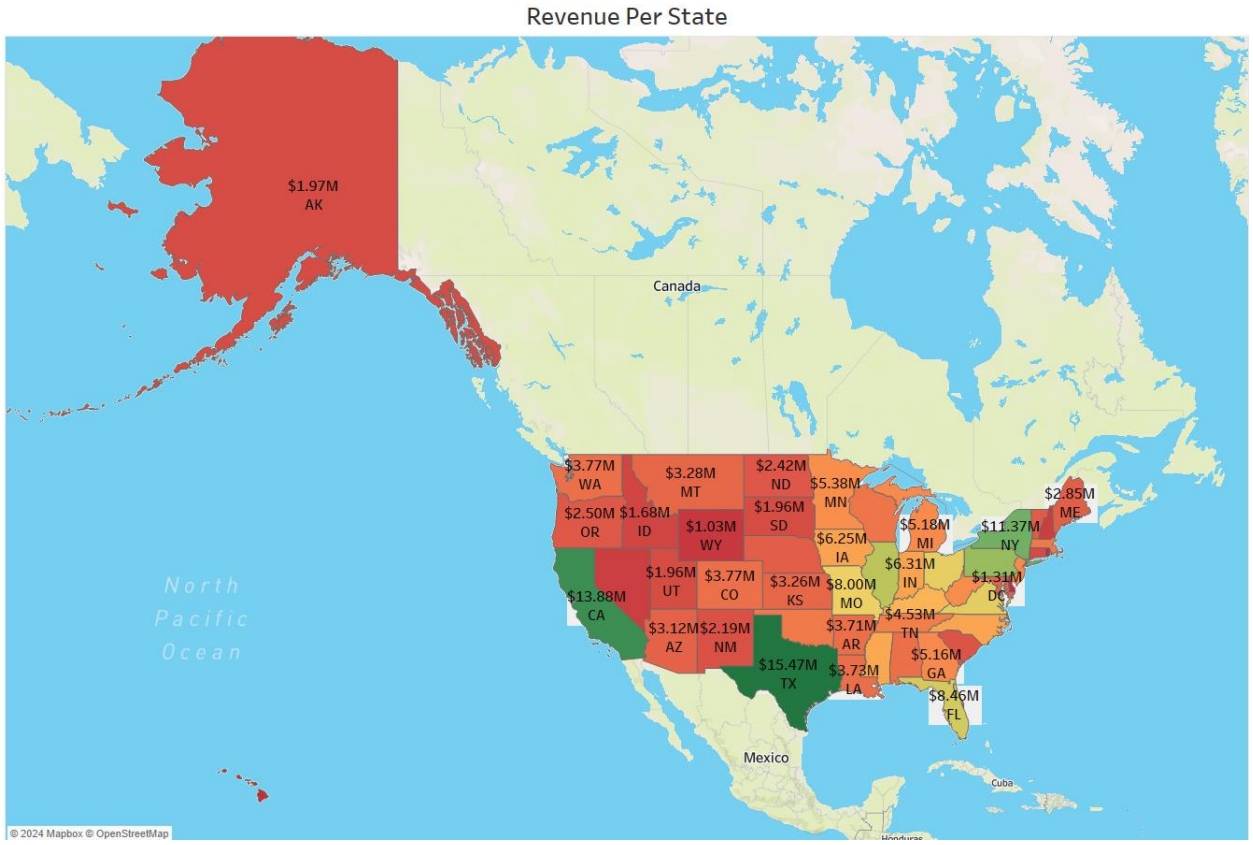
The "Quantity Discount Correlation" graph at the top center shows that correlation between discounts offered and sales generated during the discount period.

Overall, this dashboard provides a wealth of insights into the company's revenue data. It can be used to identify trends, track performance, and make data-driven decisions.

3. Chart Inferences and Description:

Graph 1: Map Visualization for Sales Distribution

Map visualization is a dynamic method of presenting spatial data, transforming geographical information into insightful and interactive displays. Utilizing mapping tools like Tableau, it allows users to represent data points on a geographical map, offering a visual understanding of location-based patterns, trends, and relationships. From pinpointing regional sales performance to identifying distribution networks, map visualizations enhance data comprehension by providing a spatial context. Users can customize maps with diverse layers, markers, and color schemes, enabling a comprehensive exploration of geographic data that aids in strategic decision-making across industries such as logistics, marketing, and public health.



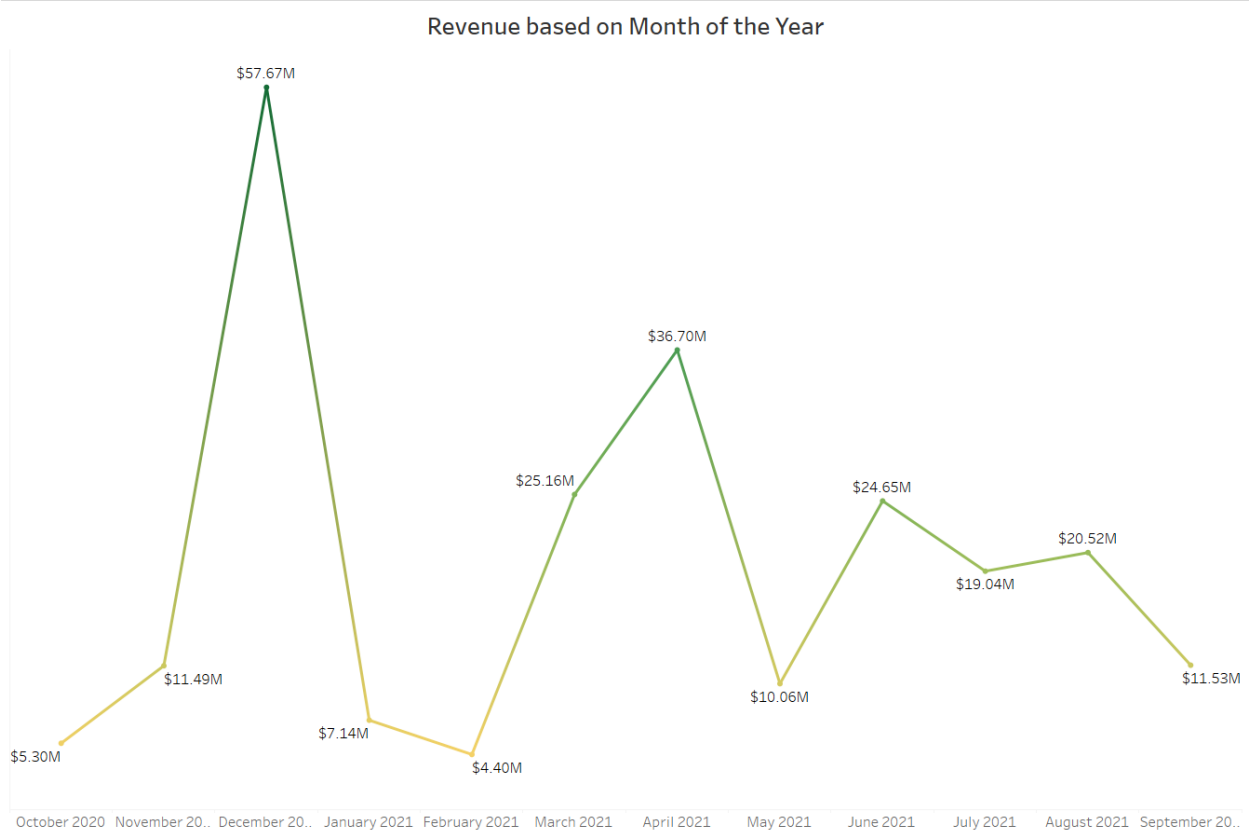
The above graph shows the sales data related to each state in the US from the highest sales (represented in Green) generated to the lowest recorded sales (represented in Red). From the graph, we see that Texas state has the highest recorded sales of $15.47 million for the year 2020 – 21 and the lowest sales is

recorded by the state of Wyoming ($1.03 million). This shows the highest and lowest spending states.

The above graph also has the option of getting to know more about the state by clicking the link once a state has been selected therefore helping the user to gain knowledge and insights on the sales number.

Graph 2: Line Graph for visualization of sales in each month

A line graph is a versatile and effective visualization tool for showcasing trends, patterns, and relationships within numerical data over time. With a straightforward representation of data points connected by lines, it allows users to easily track changes and fluctuations. Line graphs are particularly valuable for illustrating temporal sequences, enabling the identification of growth, decline, or cyclic patterns in datasets.



The above graph shows the sales done in period of one year between October 2020 and September 2021, we see a trend of highs and lows but particularly we see the highest sales in the month of December. This is likely due to holiday

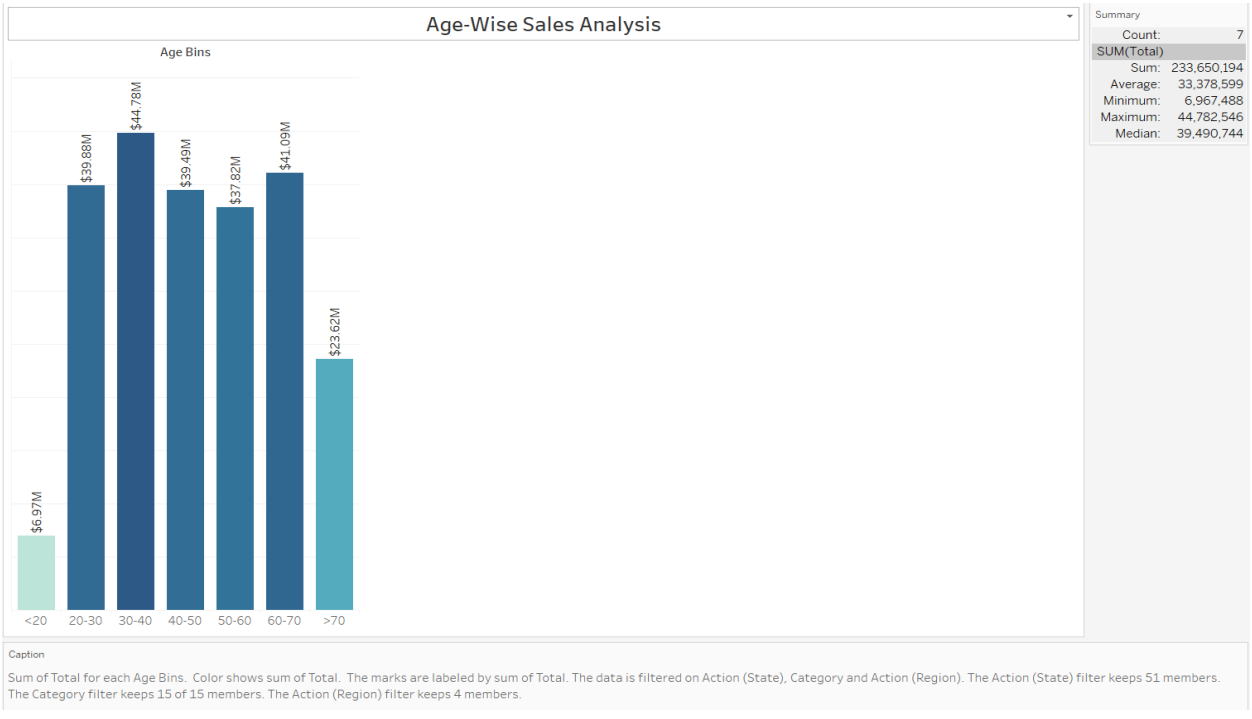
season in the US. Everyone is buying gifts and other things for family, friends and colleagues hence generating a big revenue and sales. The second biggest high is during April. The lowest sales are done in February 2021.

The graph does not consider income levels, which could be a factor influencing purchasing power across different age groups. It is possible that different age groups have different preferences for the products or services being sold. This could skew the sales figures towards certain age groups.

Revenue is relatively consistent from March to June, ranging from $19.04M to $20.52M, suggesting a period of stability.

Graph 3: Bar Chart for visualization of spending demographic

A bar chart is a concise yet powerful data visualization tool that presents categorical data with rectangular bars of varying lengths. Each bar's height correlates with the corresponding data value, providing a quick and intuitive comparison across different categories. Whether illustrating sales by product category or survey responses by option, bar charts offer a straightforward and effective means of conveying insights. Their simplicity and clarity make them a staple in data analysis, enabling users to make informed decisions based on easily interpretable visual representations.

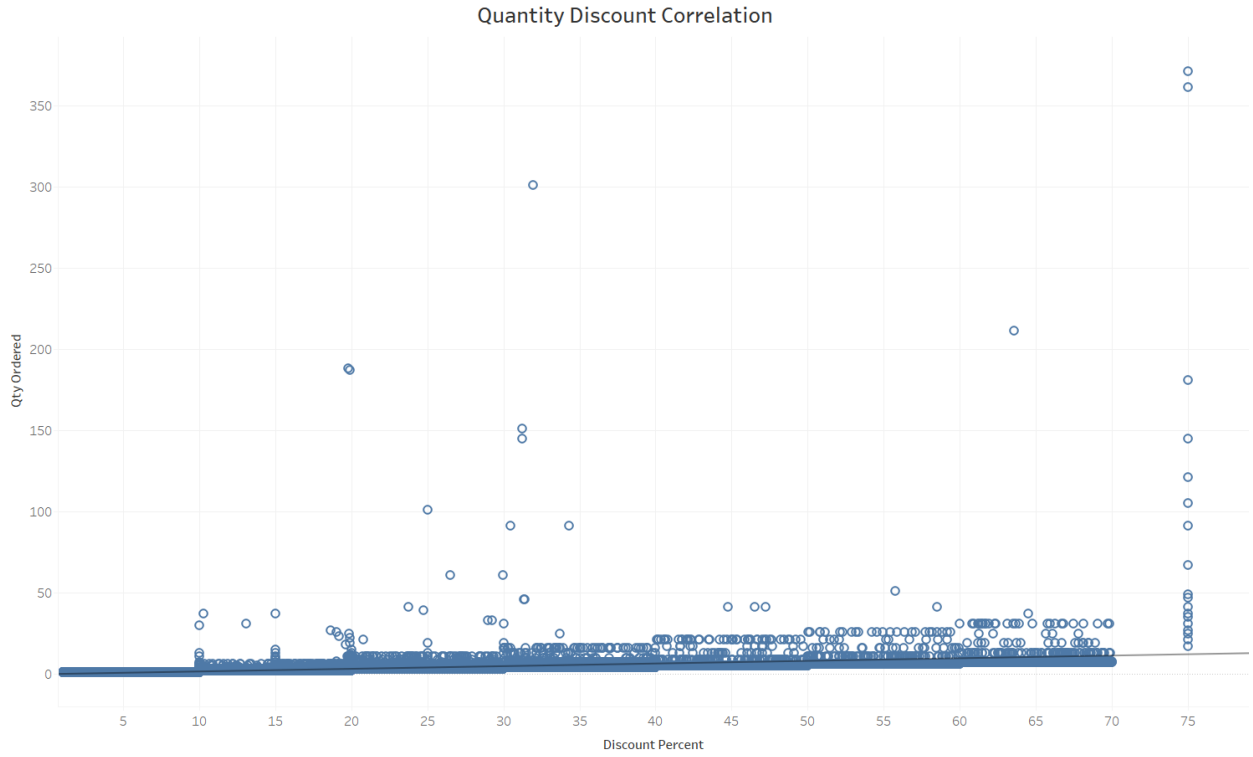


The graph shows the spending demographic i.e. the age group that spends the most or buys the product and it shows a clear picture that the age group of 30-40 years buys the most products hence contributing to the highest sales of the year (~45 million dollars). The lowest sales come from the group of demographics of <20 years old (~20 million dollars).

The average sales for the year were about $33.3 million.

Graph 4: Scatter Plot Graph to visualize Correlation between Quantity purchased and Discount

A scatter plot is a visual representation of data points on a two-dimensional graph, illustrating the relationship between two variables. Each point on the plot represents a unique observation, with the horizontal and vertical axes indicating the respective values of the variables. Scatter plots are instrumental in revealing patterns, correlations, or outliers within datasets. Their simplicity allows for a quick assessment of data distribution and the identification of potential trends, making them a valuable tool for understanding relationships and informing decision-making in various analytical scenarios.



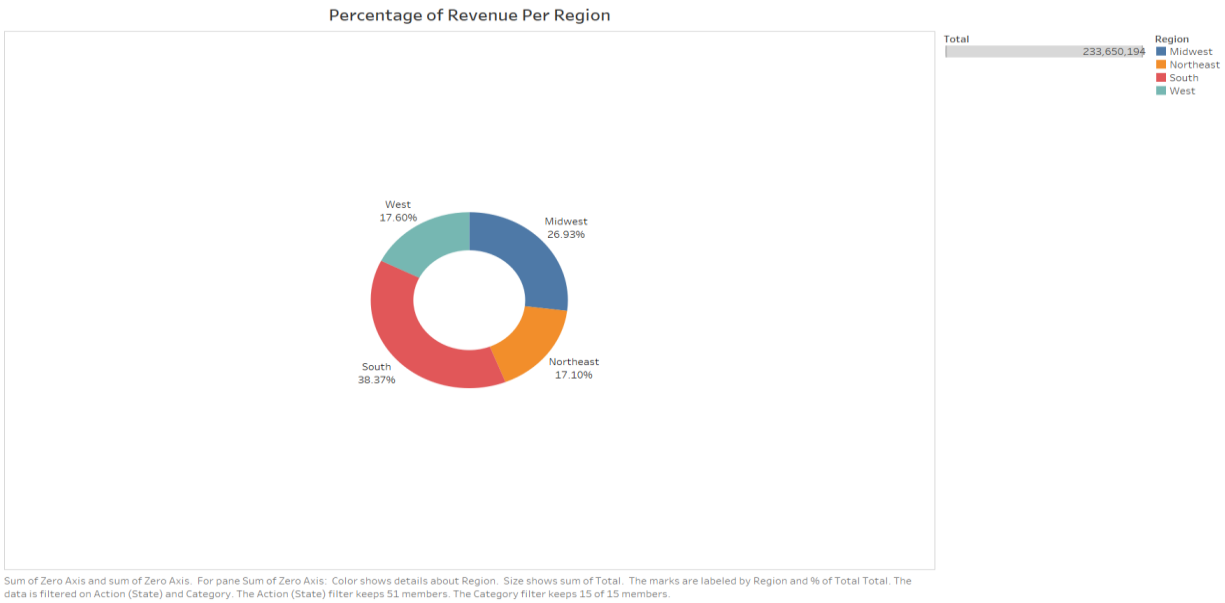
Discount Percent vs. Qty Ordered. The data is filtered on Action (State), Category and Action (Region). The Action (State) filter keeps 51 members. The Category filter keeps 15 of 15 members. The Action (Region) filter keeps 4 members. The view is filtered on Discount Percent, which includes values greater than or equal to 1.00.

The graph represents the correlation between the Quantity ordered and the Discounts offered on the products. There is clear indication that with discounts the quantity ordered is high and the trend only increases as the discount percentage increases hence positive correlation. For some bulk orders, we see a huge increase in the quantity ordered as the discount increases (the highest being 75%).

The graph does not differentiate between different product categories. It is possible that the relationship between quantity and discount varies depending on the type of product being sold. The graph could inform targeted marketing campaigns, offering larger discounts to customers who are more likely to purchase in higher quantities.

Graph 4: Donut Chart for distribution of sales from different regions of the country

A donut chart is a visually appealing and concise data visualization tool that presents categorical data in a circular format. Like a pie chart, it showcases proportions of a whole, but with a hole in the center. Each segment represents a category, and the size of the segments reflects their relative proportions. Donut charts are effective for displaying simple data relationships, emphasizing individual categories, and maintaining a clean aesthetic. While not as common as other chart types, donut charts are suitable for conveying percentages and facilitating a quick understanding of categorical distributions.

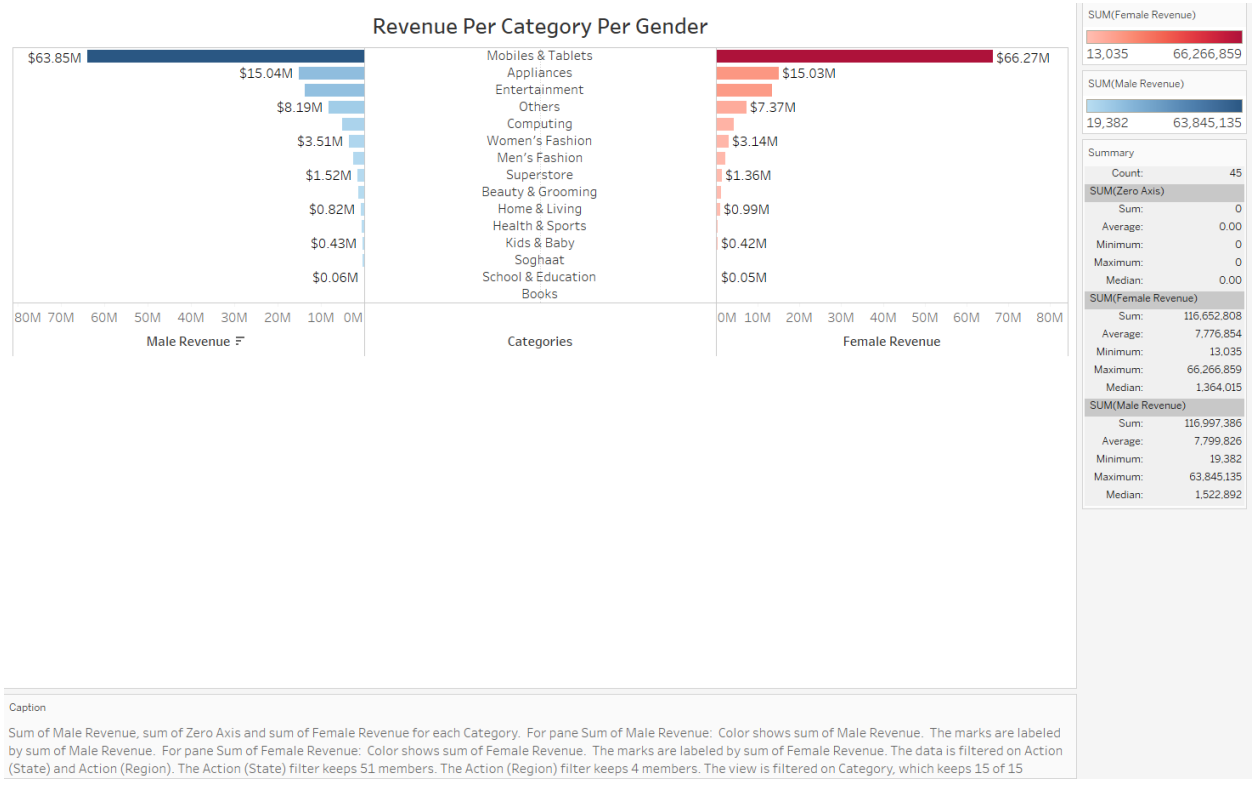


The pie chart shows that the "South" region generates the most revenue, followed by the "Midwest" and "West" regions. This suggests that the sales are done dominantly in the southern United States.

This chart shows us the distribution of sales across the region of USA.

Graph 5: Butterfly Chart for comparison of expenditure based on genders

A butterfly chart, also known as a diverging bar chart, is a unique data visualization that effectively displays the distribution and comparison of two sets of data along a central axis. The chart features two mirrored bar charts, resembling butterfly wings, where the central axis represents a common baseline. Each bar's length represents the magnitude of a variable, and the space between the bars allows for a clear comparison between the two datasets. Butterfly charts are particularly useful for illustrating the positive and negative variances between two categories or variables, making them valuable for insightful data analysis and decision-making.



Insights:

Mobiles & Tablets: This category generates the highest revenue for both genders, with males contributing $63.85M and females $66.27M.

Appliances: The revenue from this category is relatively low, with males contributing $15.04M and females $15.03M.

Entertainment: This category is interesting as it shows revenue from male consumers, amounting to $13.51M and females having $13.45M

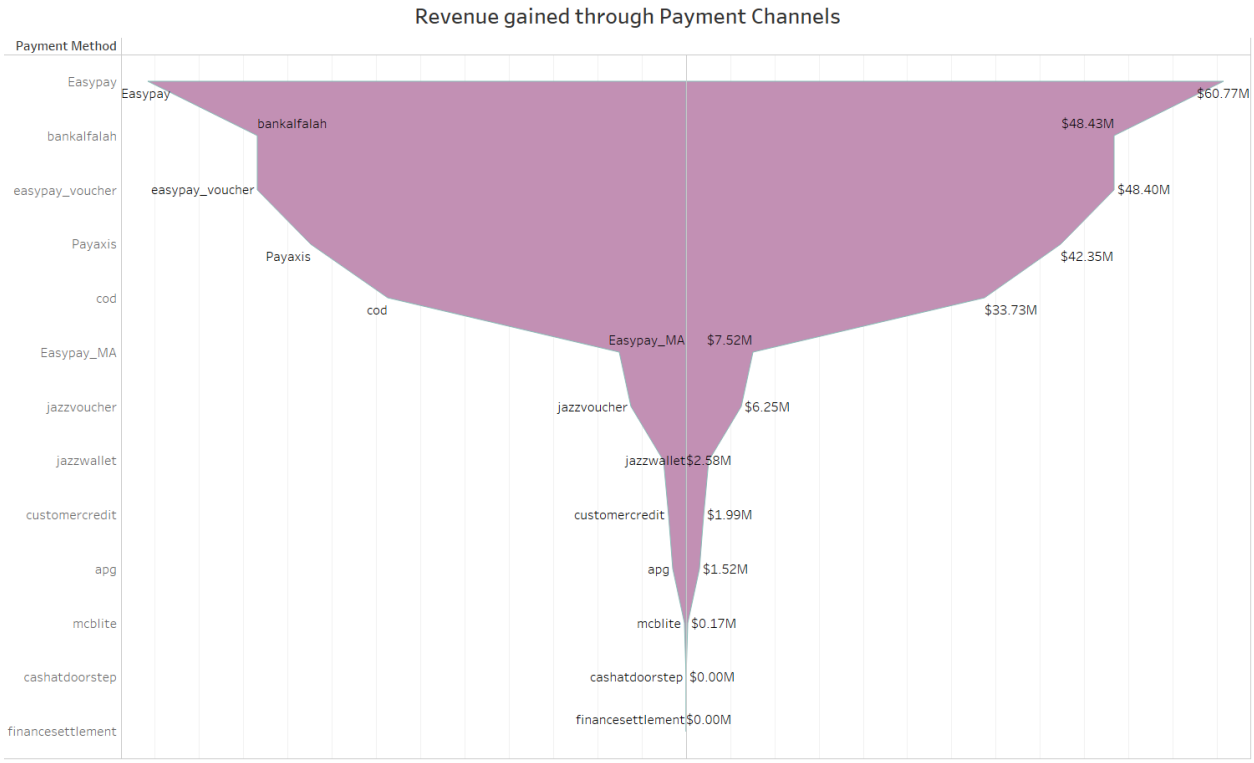
Computing: Like Entertainment, this category also revenue from male and female consumers, but the amount is lower at $0.65M.

Other Categories: The graph also includes several other categories such as Women's Fashion, Superstore, Beauty & Grooming, Home & Living, Health & Sports, Kids Baby, and School & Education.

The graph clearly shows that the "Mobiles & Tablets" category is the most profitable for both genders. There are certain categories like "Entertainment" and "Computing" where only male consumers seem to contribute to the revenue. The revenue contribution from female consumers is significantly lower than that from male consumers across all categories.

Graph 6: Funnel chart representing the income generated from various payment channels

A funnel chart is a compelling and intuitive data visualization that represents a sequential process, showcasing stages of progression from a broad start to a narrower endpoint. Resembling an inverted pyramid, the funnel chart illustrates the gradual reduction in quantity or values as they move through each stage. This visual tool is effective for tracking conversions, identifying bottlenecks, or analyzing sales pipelines. The varying widths of the funnel segments visually emphasize the impact of attrition or progression at each stage, providing a quick and insightful overview of the efficiency of a process or workflow.



Sum of Neg_Payment and sum of Value for each Payment Method. For pane Sum of Neg_Payment: The marks are labeled by Payment Method. For pane Sum of Value: The marks are labeled by sum of Value. The data is filtered on Action (State), Category and Action (Region). The Action (State) filter keeps 51 members. The Category filter keeps 15 of 15 members. The Action (Region) filter keeps 4 members.

Easypay: This payment channel generated the most revenue by far, at $60.77M. This suggests a strong preference among customers for this payment method.

Bankalfalah and Payaxis are also significant: These two channels generated $48.43M and $42.35M, respectively, indicating they are also popular choices for customers.

COD: Cash on delivery generated $33.73M, suggesting that some customers still prefer to pay upon receiving their goods.

Mobile payments: Easypay MA and jazzwallet generated $7.52M and $2.58M, respectively, showing that mobile payment methods are gaining popularity.

Wide range of payment channels offered: The company offers a variety of payment channels to cater to different customer preferences.

Revenue concentration: Most of the revenue is concentrated in a few payment channels, with Easypay accounting for over 90% of total revenue.

Regional or product-specific preferences: It is possible that payment channel preferences vary by region or product category.

Potential for optimization: The company could potentially optimize payment processing fees or offer incentives for certain channels to increase revenue or customer satisfaction.

4. Conclusion and Final Inferences

- Strong revenue performance: The company or organisation is generating significant revenue across various regions and demographics.
- Dominant payment channel: Easypay is the preferred payment method by a large margin, suggesting efficient integration and customer trust.
- Seasonal fluctuations: Revenue shows clear seasonality, with peak sales in December and a noticeable dip in February.
- Gender-specific preferences: Product categories like Mobiles & Tablets attract more female customers, while Appliances are favored by men.
- Mobile payments emerging: Mobile payment options like Easypay MA and jazzwallet are gaining traction, indicating a shift in customer behavior.
- Diversify marketing efforts: Increase focus on promoting underutilized payment channels like mobile wallets to address market shifts.
- Seasonal sales strategies: Implement targeted campaigns and inventory adjustments to capitalize on peak seasonality and mitigate off-season drops.
- Category-specific promotions: Tailor marketing and product offerings to cater to gender-specific preferences within each category.
- Regional analysis: Identify regional variations in revenue, payment preferences, and product demand to optimize marketing and resource allocation.
- Payment fee optimization: Negotiate lower processing fees with providers or consider offering incentives for preferred channels to improve margins.
- Customer feedback: Gather customer insights through surveys or focus groups to understand payment preferences and improve user experience.
- Further Exploration can be done by adjusting parameters in the dashboard such as categories and selective parameters.